Contractor Report 145378

(NASA-CR-145378) AVIONICS PERFORMANCE ANALYSIS: A HISTORICAL REVIEW AND A CURRENT ASSESSMENT OF FLIGHT INSTRUMENTATION AND CONTROL SYSTEMS IN CIVIL AVIATION Final Report (Illinois Univ., Urbana-Champaign.) N78-32094 Unclas G3/06 31573

AVIONICS PERFORMANCE ANALYSIS

A Historical Review and A Current Assessment

of

Flight Instrumentation and Control Systems

in

Civil Aviation

FINAL REPORT

Prepared for
National Aeronautics and Space Administration
Langley Research Center

Contract NAS1-15145

by

Aviation Research Laboratory
Institute of Aviation
University of Illinois at Urbana-Champaign

July 1978

TABLE OF CONTENTS

INTRODUCTION

Flight instrumentation and control systems have played an important role in the advancement of civil aviation from its early barnstorming days to its present stature as the safest form of commercial transportation. The first part of this report discusses some of the early systems which initiated this progress. While these approaches may seem humorous when compared to the sophisticated systems of today, they provided a firm foundation for later advances.

The second part of the report briefly discusses a number of the substantial contributions being made today, and projected for the future, by flight instrumentation and control systems, in the areas of safety, cost reduction, and increased capabilities. Following that, a more detailed discussion of the cost/performance considerations is given; these techniques are especially important when attempting to determine the relative values of two comparable systems or when determining the absolute worth of a system.

HISTORY OF FLIGHT INSTRUMENTATION AND CONTROL SYSTEMS

Pioneer Period

When the Wright Brothers first flew their airplane at Kitty Hawk on Dec. 17, 1903, no one could have predicted the sophistication to which aeronautics would develop in just a few short years. The airplane developed from an experimental craft barely capable of lifting its and its pilot's weight off a North Carolina sand dune to an important implement of war in twelve short years, and ten more saw its emergence as a practical method of carrying cargo, mail, and passengers over transcontinental distances at night and in foul weather. This tremendous progress is attributable to two factors: the development of more reliable airframes and engines, and to the development of instrumentation to enable the pilot to navigate his craft safely and accurately in all weather.

The lack of importance placed upon instrumentation is clearly illustrated by the example of the landmark flight of the French aviator Louis Bleriot, the first man to fly across the English Channel on July 25, 1909. Although the weather was clear when he took off from France at 4:41 a.m., a mist came up in the channel which prevented him from seeing both the shoreline and the torpedo boat sent to follow him. He had not thought to equip himself with even a pocket compass, and as a result had to let the airplane fly its own course for a time. Luckily, he was headed in the right direction and eventually encountered the English coast near Dover and landed after a flight of 37 minutes for the 31 mile distance. As will be shown, the later airmail aviators were more reluctant to trust themselves to such luck.

The period between 1909 and the First World War saw increasing numbers of important developments. Serious research into aviation was pioneered by both individuals and, increasingly, government agencies. England in 1909 created its famous Advisory Committee for Aeronautics, which did pioneer research into airframe testing. Somewhat later the United States followed suit by creating the National Advisory Committee for Aeronautics and its experimental facility at Langley, Virginia, which grew to be the largest and best equipped research facility in the world. Although tremendous advances in engine and airframe design were going on during this period, important developments were also occurring in the instrumentation area.

The first experimental use of a radio for transmitting a message from an airplane was made in August, 1910, by J. McCurdy. At an aviation meeting in Sheepshead Bay, N.Y., McCurdy transmitted a Morse code message to a receiver on the ground.

The following year saw the first airmail flights. The first demonstration was in England, where on September 9 a flight was made to carry mail between Hendon and Windsor. Soon afterward, flights were made during the week of September 23 to 30 from Nassau Boulevard to Mineola, Long Island. Although such demonstrations were looked upon as curiosities at the time, their importance cannot be underestimated in view of the fact that airmail carriers, being among the first regular commercial users of aircraft, spearheaded future developments in navigation and blind flying.

Also appearing during this period was a device which would later be developed into the 'automatic pilot'. In 1906, Lawrence Sperry began experimenting with gyroscopic aids for attitude reference and stability augmentation. Early aircraft were typically very unstable, and it normally required the pilot's complete attention to operate the controls. As longer flights became practical due to advances in reliability, navigation and other pilot duties became increasingly important, limiting the attention which the pilot could pay to the basic task of flying the airplane. It became clear that the development of a system to operate the controls to maintain a preset attitude and course was a very important task. Sperry first installed a gyroscopic instrument i.. an airplane in 1909 [10,11] and in 1914 demonstrated an automatic gyrostabilizer that won him a $10,000 prize in France's Grand Prix for Safety in Flight. This device used four gyros to give both aileron and elevator control. Further work resulted in the development of the Sperry Gyropilot, one of the first practical autopilot systems.

World War One

The First World War was, from an aviation standpoint, mos notable in that it pointed out many new applications for the airplane, and placed many new demands upon its reliability and serviceability. Early in the war, nobody regarded the airplane as a serious weapon, and its use was mainly confined to reconnaissance over enemy lines. Pilots went armed only with their service revolvers, and had great fun taking pot-shots at enemy aircraft to make sure that they kept their distance. Thus, aerial combat began, and soon airborne armament progressed through machine gun equipped planes to craft equipped with bombs for dropping on enemy munitions plants. Technological escalation set in, and soon the advent of propeller synchronized machine guns and special bomb sights turned the airplane into a serious weapon. The demands placed upon airplanes as weapons led to developments in instrumentation systems to increase their accuracy and reliability, and these will be discussed in detail later.

Transcontinental Air Mail

Consideration of the use of airplanes for carrying mail, which was begun in 1911, was suspended during the war. However, soon after the war ended in 1918, an appropriation of $100,000 was made by Congress for experimenting with mail airplanes. As a result, the first regular nonmilitary air mail service was begun on May 15, 1918 [1]. The route was from New York to Washington, and although it was initially flown by Army pilots and equipment, the service was soon taken over by the post office and its own pilots. After the reliability of this route was shown, plans were begun to establish a transcontinental route. During the day, the mail was to be flown, and at night it would be put aboard a waiting train on which it would be carried until dawn the next da' . Night flying had of course been done during the war, but this was because pilots thought the extreme risk was less than that of enemy fire. Flying in the darkness during peacetime was quite another matter, and most pilots were not prepared to undergo the risks. It was therefore felt that air mail delivery would not become fully practical until the problem of night flying was solved.

Finally an attempt was made to demonstrate the practicality of night flying. On February 22 and 23, 1921, a through trip from San Francisco to New York was scheduled. The pilot, appropriately named Jack Knight, flew over the dark leg guided by bonfires lighted by interested farmers and chambers of commerce [1]. The success of the flight prompted Congress to approve an appropriation to equip a part of the transcontinental route for regular operation at night. Emergency landing fields were set up along the way, and at these and other selected locations along the route, flashing beacons were installed. Landing lights and floodlit hangars were provided at all landing fields. The first regularly scheduled trip was made on August 22, 1923 between Cheyenne and Chicago, ushering in a new era in aircraft navigation. Later, as both mail and passenger carrier service expanded, the 'airways' defined by these beacons also increased in number. Soon radio beacons supplemented and then took over from the lighted beacons, and by 1938 the Civil Aeronautics Authority controlled over 23,000 miles of domestic airways equipped with lights or radio beacons. In spite of the depression, development of the airway system had continued at a rapid pace.

Other Early Developments

The Sperry Gyrostabilizer, described earlier, had been originally conceived as a device for holding an airplane on course in the belief that automatic stability was essential for safe flying. Airframe development soon resulted in aircraft which were sufficiently stable as not to require automatic stability augmentation. (In a later section, it will be seen that an opposite trend is developing -- auxiliary stability augmentation systems, such as active controls, are replacing the inherent airframe stability to aid energy efficiency.) However, Sperry was not to be daunted, and continued development of his apparatus to the point where it could completely take over maintaining course, altitude, and attitude of the airplane. Longer flights were becoming more common as regular passenger and mail routes were instituted, and the desirability of an automatic pilot device to help reduce pilot fatigue began to become apparent. This instrument (which will be described in more detail later) was initially very unnerving to pilots - there is a natural distrust to any mechanical device which its designers claim will do a job formerly performable only by a human. When the device performs the job more accurately and reliably than the human pilot, aggravation is added to distrust. Pilots demanded inclusion of extensive override capabilities, and these were incorporated. These considerations should be quite familiar to anyone who has had any acquaintance with discussions about current autoland systems, and perhaps the ultimate acceptance of and dependence upon autopilots should illustrate the eventual outcome of the current debate.

Another interesting development during this period was that of the radio (or radar) altimeter. The device, familiar to everyone today, was quite amazing at the time of its introduction in 1938. In early tests, the altimeter registered distance to the ground so accurately that it was possible to detect the presence of the George Washington Bridge as the airplane passed over it during the test. The value of this device as an aid to flying in fog and darkness was immediately apparent.

The same year saw the development of the first index-finger radio direction finder. Early direction finders depended upon a crew member manually rotating a directional antenna to determine the direction from the aircraft to the transmitter. The automatic direction finder performed this chore itself, allowing it to be used easily by a solitary pilot. In addition to being able to home in on beacon transmitters, the instrument can also be tuned in on any AM broadcast transmitter. These can therefore be used to navigate, or (more likely) to keep the pilot entertained while his autopilot does the flying.

EARLY FLIGHT INSTRUMENTATION

Although the earliest aircraft were only sparsely instrumented, as flights became longer the need to monitor engine performance and to navigate became apparent, and soon instrumentation began to be developed especially for aircraft applications. By the time of the First World War, the military had set rather rigid standards designating the instrumentation with which all aircraft was to be equipped. The 1915 Army Signal Corps requirements for cockpit instrumentation specified that all aircraft except trainers must carry an altimeter, airspeed indicator, compass, pressure gauges, fuel gauge, water temperature gauge, tachometer, and clock. (These are the same instruments currently required for VFR flight.) Although several of these instruments were straightforward adaptations of designs for ground use, several others had to be developed especially for use in an aircraft.

Air Speed Indicators

The accurate determination of airspeed was important to early aviators for two reasons. The first is that airspeed must be known to insure operation of the aircraft within a safe range to avoid stalling on the one hand and excessive structural loads on the aircraft on the other. The second is that speed is a necessary variable for dead reckoning navigation. A major problem, however, is the fact that these two requirements are actually calling for different types of airspeed indications. Navigation requires knowledge of the rate of progress of the aircraft through the air, independent of air density or temperature. By correcting this measurement for wind velocity, the progress of the aircraft over the ground can be calculated. Such a measurement is known as 'true airspeed'. The lift and dynamic load upon an aircraft, on the other hand, is dependent upon the density of the surrounding air, and any airspeed indicator must be able to tell the pilot the lowest and highest safe speed regardless of the condition of the air through which the aircraft is flying. Instruments were developed to measure this quantity, which came to be known as 'indicated airspeed'. Since early aircraft had very low payloads, most could not carry both types of instruments. The emphasis being not unexpectedly upon safety rather than accurate navigation, the choice was usually made to equip an aircraft with an indicated airspeed rather than a true airspeed gauge.

True Airspeed Indicators. A number of instruments were developed for measuring true airspeed, most owing their origins to anemometers used for meteorological purposes. Most of these are similar in principle; a moving flow of air is used to spin a windmill of some sort, and the speed of rotation of the windmill's shaft is some function of the airspeed. These instruments, primarily developed in Europe, were of two primary types: cup anemometers (similar to those used by weather stations) and vane anemometers (similar to ordinary windmills) [12].

A representative cup-type anemometer, the Morell, was developed in Germany and used on many aircraft there until around 1930. It consisted of cups rotating in the usual manner, with pivoted weights attached to the shaft (displaced by centripetal force) causing axial movement of an attached rod. This rod is coupled to the indicator dial. The instrument is self-contained, and is mounted on a strut, being read at a distance by the pilot. Although accurate at high speeds, it was unreliable at low speeds due primarily to sluggishness of the centrifugal mechanism.

Vane, or windmill, anemometers were developed in England, and also by the Bureau of Standards in the United States. Most anemometers of this type used some form of commutator to produce electrical pulses of a frequency proportional to airspeed, and had an circuit coupled to the indicator needle which counted the pulses received in a unit of time. The use of electrical signals enabled mounting the anemometer remotely from the indicator, which was an improvement over the Morell instrument, but the problem of inaccuracy at low speeds caused by inertia and friction in the mechanism persisted.

A third type of airspeed indicator was developed in France and Italy. The French instrument, the Eteve, was used on early types of aircraft, and embodied a flat plate mounted normal to the airflow and coupled to a spring. An indicator needle was connected directly to the plate, thus requiring it be viewed at a distance by the pilot. Accuracy of this instrument was low, and it was not even calibrated in units of speed. The only indications used were for minimum, cruise, and maximum speed.

Pitot-Static Indicators. The pitot-static instrument is an adaptation of a method used by Pitot in 1732 to measure the speed of a river. Since these instruments are in common use today, only the briefest explanation is in order. Two tubes are used, one open ended and the other - the static tube - is closed at the end and is mounted nearby. The static tube has several small holes in the side near the end, thus exposing it to atmospheric pressure in the vicinity of the open tube. The pitot tube leads to one side of a diaphragm, and the static tube to the other. The difference in pressure, proportional to the square of the wind velocity, causes a deflection of the diaphragm. The center of the diaphragm is connected in one of a variety of manners to an indicator needle. It is obvious that differences in air temperature and density will cause the indicated airspeed to vary from the true value, but the indicated value will accurately reflect the effect of the airflow on the control surfaces. Since this information is crucial to control of the aircraft, it was considered more important than the true airspeed value, and this led to the pitot-static instrument supplanting other types of airspeed indicators.

Much research was done toward relating indicated airspeed to true airspeed. As a result, tables and slide-rule type calculators were made available to allow determination of true airspeed from the indicated airspeed, air pressure, and air temperature values. Modern transport aircraft carry analog air-data computers capable, among other things, of performing this calculation.

Tachometers

The tachometer, or engine revolution indicator as it was known early in the century, is used to indicate the revolution rate of the engine's crankshaft. This was thought to be the single most reliable indicator of the condition of the engine for several reasons. First, it indicated directly any falling off in speed (thus power) of the engine. Second, it facilitated accurate control of the engine to avoid 'overrevving' it. Third, it could be used during 'running up' to determine that the engine was ready for flight. Many types of tachometers were developed for early aircraft, but only a few will be examined here.

Centrifugal Tachometers. This type of tachometer uses a Watt-type governor attached to a shaft coupled to the crankshaft of the engine. It is simple in principle: the spring-loaded weights are displaced from their rest position in proportion to the speed of rotation of the shaft. This displacement is used to drive an indicator needle, displaying the engine speed. Accuracy was quite good, typically within 20 RPM between 600 and 2600 RPM [12]. The main problem with this type of instrument involves coupling the weights to the indicator. If they are coupled directly, the indicator must be situated near the engine - clearly a disadvantage for large or multi-engined craft. If they are coupled through a flexible shaft to a remote indicator, the shaft is a likely source of trouble.

Chronometric Tachometers. This type of indicator operates by mechanically counting the number of revolutions performed during a time interval of known duration. Examples of this type of instrument were the French Jaegar, German Bruhn, and American Van Sicklen tachometers. Although the actual mechanical operation of such an instrument is very involved, its basic form consists simply of resetting an indicator, advancing it based on the number of revolutions of the input shaft, displaying it to the pilot after a fixed period of time (as determined by a clockworks mechanism which is a part of the instrument), calibrated in revolutions per minute, and starting the operation again for the next time period. To improve the performance, one part of the instrument is determining the next reading while another identical part is being initialized for the next operation. However, this causes any changes in the reading to occur in jerks during the switchover from one section of the instrument to the other.

Electric Tachometers. The development of the electric tachometer is in several ways indicative of the whole direction of instrumentation design during the past forty years. Early electric tachometers were similar in principle to those of today -- either a commutator attached to the engine or the ignition system was used to charge a capacitor, which is discharged through a resistor. Either the average capacitor voltage or discharge current is measured, the result being proportional to engine speed. The problems with electrical tachometers were just the opposite of those expected today; they tended to be inaccurate (due to temperature and humidity variations), more difficult to maintain, less robust, and more bulky than their mechanical counterparts. It is a testament to the development of electronic devices that this situation has reversed itself to the extent that it is difficult to imagine these complaints ever having merit!

Orientation Aids

Instruments in this category include such things as inclinometers, attitude indicators, and the various gyroscopic instruments performing similar functions. Early in the history of aviation, it was found that even the most experienced pilot has extreme difficulty maintaining control of his aircraft in blind-flying conditions without the aid of instruments. Thus, it was impossible to use aircraft in any situation requiring maintenance of scheduled operations under unfavorable weather conditions before instruments providing orientation information were available.

Instruments in this category aid blind flying by providing two types of information. First, they indicate small disturbances from the steady state, so that the pilot may react to keep the disturbances as small as possible. Second, they indicate the aircraft's attitude during steady state flight. Early instruments generally worked on the principle of the spirit level, while later instruments typically used gyroscopes.

Early Inclinometers. With the exception of the kymograph, which used the shadow of the sun to determine the aircraft's orientation (making it of dubious value in a blind-flying environment!) most early inclinometers were one of two types. One was the pendulum inclinometer, comprised of a damped pendulum arm attached to an indicator needle, such as that marketed by Sperry. The other, and more common, was the bubble inclinometer and the similar ball in tube type.

Most bubble inclinometers used a glass tube filled with a mixture of water and glycerine. At one end of the tube was a small electric lamp to permit night use. Early levels had a tendency to burst at high temperatures, so later models incorporated a liquid trap at one end of the tube. Normally inclinometers were positioned across the cockpit to facilitate proper banking of turns, but early large flying boats were sometimes equipped with fore-and-aft levels to permit accurate pitch corrections during landing.

Naturally inclinometers such as these cannot give an indication of absolute vertical since the bubble position is determined by the resultant of all forces acting upon it. Although they are useless as artificial horizon indicators, they do provide an indication of whether or not the aircraft is properly banked in a turn. Under the influence of gravity alone, the bubble will move to the high side of the aircraft. However, in a correctly banked turn, the resultant force due to gravity and centripetal force is at right angles to the wingspan, and the bubble will remain centered. This permits the bubble inclinometer to be used as a turn coordinator, and in fact it continues to be used for this purpose to this day.

Gyroscopic Inclinometers (Artificial Horizons). For blind flying where it is necessary to have a vertical reference independent of the ground, it is necessary to use the equivalent of a gyroscopic device. Such a reference can provide both bank and pitch indications, and much work has gone into development of workable devices to provide it.

The gyroscopic principle is quite familiar to most: a spinning body will maintain a fixed axis of revolution if not subjected to force couples about the ends of the axis. If a force couple is applied normal to the axis of revolution, the axis will precess towards the axis of the couple at a rate proportional to the magnitude of the force and inversely proportional to the angular momentum of the gyro.

Normally the gyro is mounted in a gimbal system to allow it to maintain its axis of revolution as the aircraft moves around it. If the gimbal system were free from friction and balanced about its axes, the axis of revolution would define a fixed direction in space. There are two problems with this. First, the gimbals are not friction free or balanced, and second, the rotation of the earth causes the gyro axis to depart from vertical. It was soon realized that to overcome these problems it was necessary to deliberately impose an external force couple to maintain the accuracy of the vertical axis. In principle, a spinning top is similar to such a regulatory force, gravity and the support of the floor acting as a couple which tends to erect the top as it tips over.

All gravity controlled gyro verticals of this type have in common the difficulty that in turns, when centrifugal force causes the couple on the axis to deviate from vertical, the gyroscope wanders from a true vertical orientation. Instruments were designed to overcome this, most being in principle similar to the gyroscopic pendulum designed by Lanchester [12]. This device replaces the simple pendulum producing the erecting force couple. It consists of a pendulum mounted with its axis parallel to the direction of motion of the aircraft. On the pendulum is mounted a small rotor whose axis is oriented in a horizontal plane across the aircraft. If such a pendulum is mounted in a steadily turning body, it will rest in a position such that the resultant couple about the pendulum axis is sufficient to precess the rotor axis about the vertical axis at the speed at which the body is turning. Such a pendulum can be used as a datum towards which a gyroscope may be made to precess, allowing it to define a true vertical reference in either straight or turning flight.

Practical artificial horizons use vertical reference pendula somewhat less complex than this. For example, the Sperry-Horizon, typical of pre-WWII instruments, used a system of four pendulous vanes to maintain its erection. The vanes, suspended from the underside of the gyro housing, each partially covered an air port in the housing. If the gyro tended to depart from vertical during level flight, gravity holds the vanes vertical, and one vane closes one port, while the opposite vane, drawn down by gravity, opens its port, whereupon air is exhausted through the port, causing a force righting the gyro. The rate at which the gyro is righted, however, is so slow that forces caused by pendulum movements in rough air do not appreciably displace the gyro. Although a steady turn will displace the gyro due to centrifugal force acting on the pendula, the magnitude of this inaccuracy is small (less than 5 degrees) and the gyro corrects itself when the aircraft returns to level flight.

Automatic Pilots

As was mentioned earlier, the advent of large multi-engined aircraft, the increasing length of flights, and the necessity for the flight crews to navigate with increasing accuracy placed a larger workload upon the crew.

Automatic devices capable of flying the aircraft on a preset course progressed from being a convenience to being a necessity. Stability augmentation devices had been in existence from the earliest days of aviation, and as the need for a full blown automatic pilot increased, these obsolete devices were dusted off and developed into practical systems.

Probably the best known early autopilot was the Sperry Gyropilot. This system was based upon the artificial horizon and directional gyro instruments already in use at the time. After a course was set, the outputs of the gyro instruments were used as inputs to control servos operating the controls of the aircraft. Rather than use a mechanical linkage between the gyro instruments and the servos, which could disturb the gyros, deviations in the gyro positions are used to control air valves arranged to deflect a diaphragm in proportion to the magnitude of the deviation. The diaphragms in turn control balanced oil valves connected to pistons which control the servos.

The Sperry device was capable of controlling the aircraft on a straight course. It could either hold a preset altitude, or could hold a preset pitch to allow steady climb or descent. Small course changes could be accomplished by manipulating trim controls on the front panel of the unit. Resetting the directional gyro, as was required every 15 minutes or so, required either disengaging the autopilot or locking the rudder control.

Directional corrections were made solely by using the rudder; no attempt was made to coordinate turns. It was also necessary to have the airplane as near as possible in trim before engaging the autopilot. The reasons for this were to avoid servo oscillations caused by the system attempting to apply a control continuously against an out of trim condition, and to avoid the embarrassing situation of having the aircraft pitch or bank severely when the autopilot is released.

RADIO NAVIGATION AIDS

Early navigation was almost totally an art. The early pilot was a master at using his senses to determine direction, distance, and position. This was not difficult in the early barnstorming days when aircraft were slow, landmarks were plentiful, flights were short, and when flying was only done in good weather. However, when aircraft started to venture out of sight of land, due to the wish to fly over water and the necessity of being able to fly at night and in bad weather, seat of the pants navigation techniques had to make way for more sophisticated methods.

In the days before radio technology was fully developed, navigation aids were extremely primitive in nature. Early pilots had little more than compasses and dead reckoning calculations to help them navigate. During the early days of the transcontinental airmail service, it was necessary for pilots to depend upon farmers to light bonfires to aid them in following their route. Even when these routes were eventually developed into the first airways, the primary navigation aid was a system of light beacons set up along them. Although this aided the pilot in maintaining course during night flying, the problem of accurate navigation in bad weather and over bodies of water had yet to be solved.

There were other reasons for the development of radio navigation aids. In the early days of air transport flying, the aircraft were slow and the airways were relatively uncrowded; midair collisions were infrequent. As air traffic increased, it became necessary to maintain adequate separation between aircraft in addition to just enabling it to find its way to its destination. This placed a demand upon the pilot to know his position much more accurately than before. Early air traffic control was performed by receiving positional data from each pilot and plotting his course on a map. Detection of potential collisions depended upon the accuracy with which a pilot reported his position. Until the deployment of ground based radar tracking equipment, the inaccuracy of pilot reported positions severely limited the number of air operations which could be flown.

The increasing demands placed upon air transport also led to the development of ever larger aircraft. Naturally, as aircraft got larger, it became more and more expensive to keep them in the air. In the late twenties, it was considered sufficient to navigate to within fifteen minutes flying time of the destination. Current large jets consume fuel at the rate of around $1000 per hour, making such inaccuracies expensive on this basis alone. Thus economic pressures also indicated the need for increasingly accurate navigation equipment.

With the building of large airports in the thirties, a new problem arose -- the need for navigational aids for the terminal area. This problem has two aspects -- that of maintaining adequate separation between aircraft on the approach to insure safe operations, and that of enabling landing in marginal weather conditions such that schedules can be maintained in all but the most inclement weather.

The period between the late twenties and the end of WWII saw enormous breakthroughs in navigation systems. After the perfection of radio in the twenties, it was not long until its application to navigation became apparent. Soon many systems for both long and short range navigation became operational, and in fact the problem became more one of standardization than of availability of systems. The impact of navigation capability upon aviation was enormous; it can be easily argued that the availability of radio navigation was the single most important contributing factor to the development of the air transport industry.

Enroute Navigation Aids

One of the earliest and most common of all radio navigation aids was the Radio Direction Finder. Typically this made use of transmitted signals in th mid-frequency band (100 - 2000 kHz), hence the name MF/DF. Systems were also developed to use the very high frequency band (VHF/DF), but saw limited use, mostly by the military or in emergency situations. MF/DF is simple in principle, using a directional antenna with a sharp null lobe to indicate the direction to the transmitting station. Practical systems used two directional loop antennas mounted at right angles, called a Bellini-Tosi aerial. The loops had field coils connected across them, and a third coil (called a goniometer) was mounted so that it could rotate in the fields of the other coils through 360 degrees. Radio signals cause currents to flow through the field coils which are a function of the direction of the transmitting source. The goniometer is rotated so it can determine the proportion of current flowing in each field coil; when it is oriented properly to the azimuth of the source the induced fields cancel and no current flows in the goniometer. This 'null' position is quite critical, and often the direction of the source can be determined to within two degrees.

MF/DF systems were deployed in the late twenties, and used ground based receivers to fix the position of the aircraft. The azimuth reading was passed to the pilot via his voice radio. This system proved expensive and cumbersome, and was replaced in the early thirties by a system of nondirectional beacon transmitters (NDB's) on the ground, and an automatic direction finder reciever (ADF) on the aircraft.

A problem with ADF navigation is that, when flying directly towards an NDB (which, especially on instrument approaches, was the usual case) any crossing wind component results in flying a spiral track towards the NDB. The wind blows the aircraft off course, and since correction involves turning directly towards the NDB, the drift angle is never compensated for.

MF Radio Range (Four Course Range)

This device consists of a beacon transmitter, signals from which are beamed in (usually) four directions (Figure 1). Four course ranges were installed along airways starting in the late thirties By following the beams, it is possible to fly accurately along an airway regardless of crosswinds. Beams were situated so they pointed along airways, and additionally towards nearby runways. No beam is radiated upwards, so there is a 'cone of silence' over the beacon providing the pilot with an accurate indication of when he is over the beacon. It is thus possible to follow the beams along airways, and then to turn onto the final approach to the airport using only the radio range for guidance.

The 'beams' are produced by means of a set of directional antennas radiating Morse code signals. The antennas are arranged so a Morse code letter N is radiated on two lobes, one into each N Quadrant of Figure 1. Additionally, the letter A is radiated on two lobes, one in each A Quadrant. The A and N signals are synchronized such that when the aircraft is flying along the line dividing the quadrants, a continuous tone is heard. When the flight path deviates to either side, either the A or the N predominates. The pilot, knowing which leg he is flying, can correct his flight path by turning in the appropriate direction until he is back on course.

If there was ever any doubt about which quadrant the aircraft was flying in, there was a rather lengthy procedure to identify the aircraft's location. Another drawback with the four course range was that its accuracy to some extent depended upon the hearing acuity of the pilot. Early aircraft were not notably quiet, and this was sometimes a problem. This was aggravated by the fact that, to avoid collision with aircraft flying the beam in the opposite direction, it was customary to fly slightly to the right side of the beam. If the pilot was incapable of detecting the slight variation in amplitude when he was slightly off the beam, potentially dangerous situations could arise. The fact that the system operated in the MF band made it quite susceptible to static and interference. Thunderstorms could entirely blot out the signal,

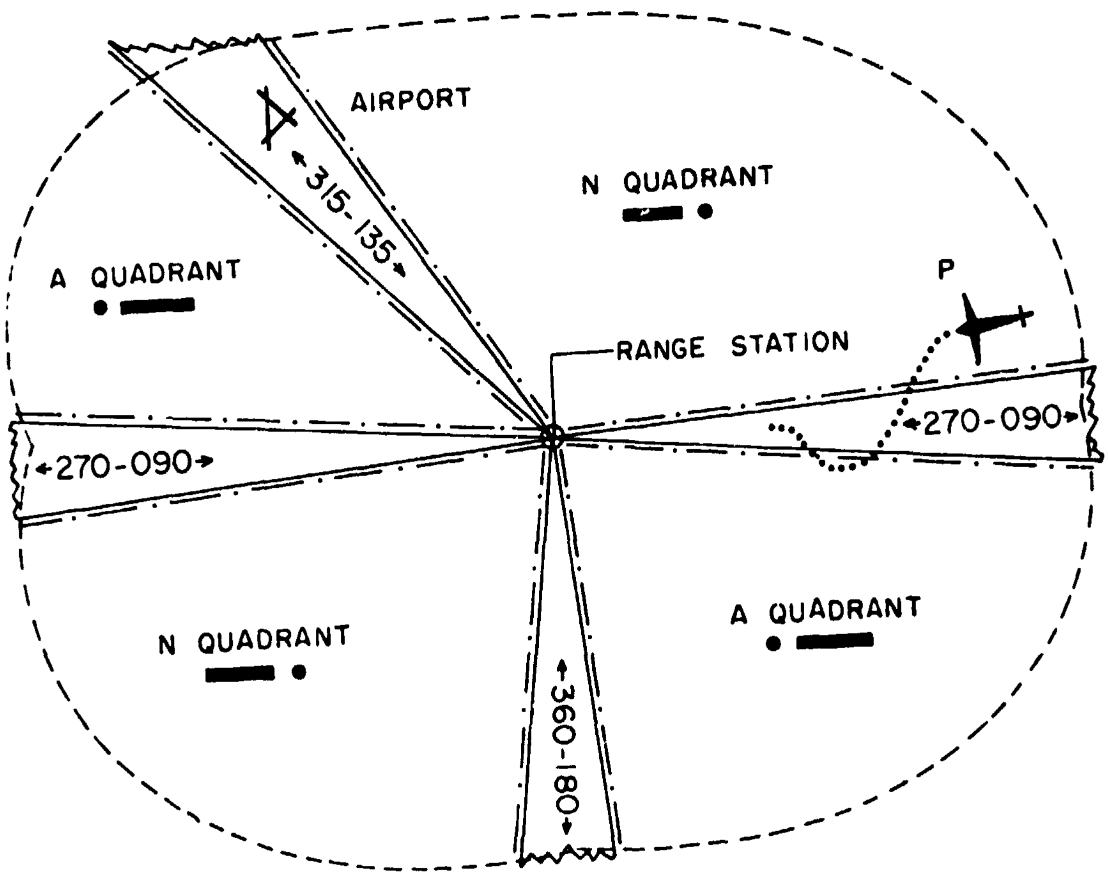

Figure 1 -- Four Course Range

making the system useless at the time when it was most needed. Finally, the limitation imposed by the small number of beams limits flexibility in selecting courses.

Consol and VOR

Several improvements on the beam-radiating principle of the four course range have been made. One of these, developed by Lorenz in Germany during WWII, was intended to guide long range bombers and U-boats. The allies were so impressed by the system that they refrained from jamming it, instead using it to guide

Allied bombing missions over Germany! Consol is similar to the MF Radio Range in that a directional array radiates dots and dashes in a multi-lobed pattern, but differs in that it has many lobes and thus defines more bearings. Accuracy is high, with bearings obtainable to within a degree, and its range was extremely long, on the order of 1500 miles. It was extensively used after the war in Europe, and in 1954 the ICAO recommended that it should be used as the interim navigational aid (along with Loran) until a more up-to-date system was developed.

The wish to have more flexibility in course selection than with the above systems, and to overcome the susceptibility to interference inherent in MF systems, led to the development of the VOR (VHF Omni-Range) [8]. This system was adopted as the U.S. standard in 1946. It had the advantage of defining an infinite number of courses (radials), and could also be used to obtain accurate position fixes by measuring the directions to more than one station. Since the readout was visual, using a meter instead of Morse code signals, it did not depend upon the hearing capability of the pilot, and also allowed normal voice radio communication to take place while using the signal to navigate. While VOR has some problems with accuracy caused by propagation disturbances, it has been extremely successful in practice.

Beam Approach and Landing Aids

The most critical phase of a flight is without question the landing. Accuracies of a few yards are not enough for landing guidance; they must be to within a very few feet. With the sluggish response of large transport aircraft, the pilot has very little opportunity to correct for navigation errors on the approach after coming out of the clouds at an altitude of one or two hundred feet. The advent of autoland systems place an even more stringent demand upon a landing guidance system -- it must be capable of guiding the aircraft through landing and rollout with no intervention from the pilot. The need for accuracy required development of systems fundamentally different from those used for enroute navigation.

The Lorenz Beam Approach System. The most famous of the early beam approach systems was that developed by Lorenz in Germany in the early 1930's [7]. Like most later systems, it used continuous wave techniques. The Lorenz system transmitted on the 30-40 MHz band. It was similar in principle to the four-course range; two highly directional antennas radiated signals whose lobes radiated outward from the runway, with the line bisecting the lobes lying exactly along the runway (Figure 2). Each signal carried a different keying code, synchronized such that when the aircraft was centered on the approach the pilot heard a steady tone. After the pilot used his ADF to roughly position the aircraft on the approach, he flew the beam down, correcting in response to which keying pattern (dots or dashes) predominated.

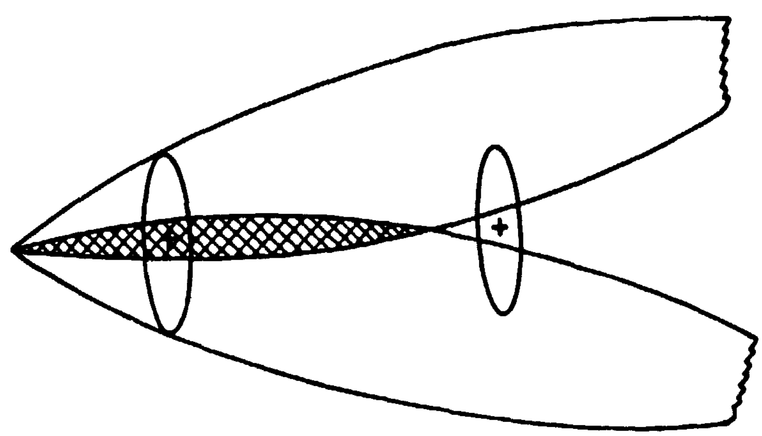

Figure 2 -- Lorenz Beam Approach System

Standard Beam Approach. An improved version of the Lorenz system was developed by Standard Telephones and Cables, Ltd in Britain. The system, called the Standard Beam Approach, attempted to overcome the fact that the Lorenz system provided no vertical guidance information. In addition to the beam signals, SBA employed two vertically radiating marker beacons, which actuated two panel mounted indicator lights in addition to producing an audible signal. This enabled the pilot to check his altitude at locations of known distance from the runway. Also, the main beam information was optionally presented by means of a meter indicating the distance the aircraft was left or right of the runway centerline.

Instrument Low-Approach System (ILS). The current ILS system is a direct descendant of the Lorenz and SBA systems. Developed during WWII, ILS differs from its predecessors in that the beam signals are tone modulated rather than using dots and dashes. Also, the system operates on the VHF band.

The principle improvement over previous systems is the incorporation of a glide slope transmitter. The operation of the glide slope system is similar to that of the lateral beam (localizer), consisting of two tone modulated signal lobes centered on a plane roughly 3 degrees above horizontal. The receivers in the aircraft compare the amplitudes of these two tones, producing a meter reading deviating from center proportionally to how far the aircraft is off the approach. Two (sometimes three) marker beacons define important locations on the approach: the outer marker is about 8 km from touchdown, the middle marker at about 1200 m, and optionally an inner marker is located about 30 m from touchdown. The outer marker is usually co-located with an NDB allowing the pilot to guide the aircraft on the approach. The middle marker defines the point at which the pilot must be able to see the ground on Category I approaches, and the inner marker defines the location at which ground contact must be established for a Category II approach.

The marker beacons provide indications to the pilot similar to those of SBA beacons, and path deviations are displayed on meters. The localizer is displayed on a vertical needle, and the glide slope is shown on a horizontal one. Although the system is subject to multipath propagation disturbances which can cause inaccuracies, in practice these are negligible and can often be cured by special site preparation. Category I approaches can be flown using only these indicators. At specially certified airports, Category II approaches can be flown down to a decision height of 100 feet, but it is required that the aircraft be equipped with special altitude measuring devices and either an autopilot or dual flight directors. Autoland systems have been designed using ILS for guidance, but at this point the inaccuracies introduced by propagation errors and the fact that the glide slope is not usable at under 75 feet prevents using ILS for automatic landings at all but a few specially certified airports.

THE CONTRIBUTIONS OF FLIGHT INSTRUMENTATION

The previous discussion has highlighted the impressive advances made in flight instrumentation and control in the past half century. These advances have paralleled the progress of electronics, from the basic triode amplifier to very large scale integrated circuits. In fact, with the exception of the replacement of reciprocating engines with jet turbine engines, most other changes in airplanes have been evolutionary, rather than revolutionary, refining past concepts to provide better operation. The major advances have been in the flight instrumentation and control areas.

These advances can be classified into three different, and sometimes overlapping, categories -- safety, cost reduction, and additional capability, to be discussed in greater detail below. The first category is one of the most important, since it has produced a transport aviation segment which has a significantly better safety record than any other form of transportation. In addition to the obvious contributions of the electronically-based modern air traffic control system, flight data and cockpit voice recorders have proved invaluable in discovering and correcting problems causing the few accidents which have occurred.

Flight instrumentation and control systems have produced substantial reductions in the cost of operating aircraft, by improving crew member efficiency, easing maintenance costs, and increasing energy use efficiency by reducing route lengths and delay. However, from a traveler's point of view, possibly the greatest advance (with the exception of safety) has been in the enhancement of the capabilities of the basic airframe, particularly in the area of all weather operations. This has changed aviation from a "barnstorming" novelty to a scheduled transportation system in which even minor delays can cause great inconvenience. Without the advances in flight instrumentation and control systems, particularly in avionics systems such as navigation or automatic flight control, none of this would be possible.

Providing Safety to Flight

While many areas of flight instrumentation and control affect the safety of a flight, only those whose primary purpose is an improvement in safety will be considered here. Those areas which have indirect, though significant, effects will be considered later. For example, modern navigation systems allow a pilot to fly from one airport to another without reference to landmarks on the ground. While this dramatically improves the flight's safety when operating in clouds, this navigational capability will not be considered as primarily a contribution to flight safety, but to expanded capabilities. In general, it will be assumed that the pilot is operating his aircraft in conditions (poor weather, night, etc.) which are appropriate for the equipment on board the aircraft.

Air Traffic Control System. The most significant contribution to air safety provided by flight instrumentation systems is, of course, the collision avoidance capabilities necessary for non-visual flight in potentially congested airspace. This is the primary responsiblity of the air traffic control system (ATC), and was initially carried out by correlating reported positions of participating aircraft at a central location, and by providing "clearances" within blocks of airspace to insure sufficient separation. While ATC is not usually regarded as a form of flight instrumentation, it is increasingly becoming an independent, redundant pilot input for position and altitude, increasing confidence in the more conventional flight instrumentation systems.

All that is needed for an aircraft to participate in this service, in addition to the equipment necessary for navigation and flight control, is a radio for communications with ATC. While in the early days of flying such radios were found only on commercial flights, today it is rare to find a plane which does not have at least one two-way radio. In the future, low cost digital data links may be used to reduce transmission errors and improve the pilot-controller interface.

Radar Separation. The advent of radar substantially altered the procedures used for collision avoidance, by allowing the controller on the ground to independently monitor the positions of the various aircraft. In addition to preventing difficulties caused by a pilot missing a position report or incorrectly reporting his position, the increased accuracy in determining the position of one aircraft with respect to another allows a substantial decrease in the separation between two aircraft (from as low as 3 nm using radar to as much as 30 nm without radar [4]), allowing more aircraft to operate in a given area.

As the number of aircraft being served by ATC radar has increased (primarily general aviation operations), the need to accurately and uniquely identify a given aircraft has become more important. This has caused the introduction of radar transponders (ATCRBS, the air traffic control radar beacon system), which can provide a unique code for every aircraft in a given sector (4096 total codes), based on a four digit number specified by the controller and selected by the pilot. The transponder can also be used to downlink the aircraft's current altitude (the beginning of a digital datalink system). Besides inproving the presentation of data for the ground controller, computer monitoring of a flight progress is possible. If the aircraft is in danger of entering another aircraft's airspace, or if its altitude may be too low for the surrounding terrain or obstructions, the controller is automatically alerted, and a warning can be forwarded to the pilot. Proposed modifications to the basic radar beacon system would allow the automatic transmission of this information to an aircraft [5], providing redundancy to the controller's instructions.

The present trend is toward upgrading both airborne and ground-based systems to upgraded third-generation air traffic control (UG3RD) capabilities with emphasis on accuracy and more automated monitoring and predicting of traffic situations. Radar digitizers and narrow-band information transmission will increase radar coverage, especially in the areas which form the borders between adjacent air route traffic control centers' airspace, as will the use of transponder-only radar in areas (such as in the mountains) where conventional radar was not usable.

Independent Collision Avoidance. Other systems have been proposed which would provide collision avoidance without the use of ground-based controllers or equipment [6]. However, these require that all aircraft be equipped with the special equipment, since any aircraft receiving a coded signal must be able to respond with its altitude and other information necessary to determine its position. Recently, systems have been proposed which use the standard ATCRBS transponder for the basic equipment, with additional equipment required only for those aircraft which desire collision avoidance capability in addition to that provided by the ATC system [9]. This provides an independent backup in case of a human or equipment failure in the ATC system. As other, more accurate navigation systems, such as GPS/Navstar, become available, systems which cooperatively transmit their position and receive the position of surrounding aircraft will be introduced, providing further redundant information for flight safety.

Ground Proximity Warning. Because of a number of serious airplane accidents, where due to distractions or misinterpretation of altitude information airplanes were flown at altitudes which cause collision with the ground, the FAA has required the installation of special electronic systems (GPWS, ground proximity warning systems) which monitor the airplane's altitude above the terrain, and its rate of descent, to provide an alarm if it detects a problem. Although the initial systems contained problems which caused the inadvertent activation of the alarm when no problem existed, the reliability of GPWS has greatly improved, and when combined with the automatic altitude warning capability provided by ATC radar in all major terminal areas, makes the occurrence of ground collisions highly unlikely. This is a relatively new requirement for commercial aircraft which reinforces the trends of more accurate spacial position determination (in this case in the vertical dimension) and the use of independent, redundant systems for added safety.

Accident Investigation. Onboard instrumentation recording important data about each flight and all conversations and other sounds in the cockpit has allowed accident investigators to determine accurately the cause of most scheduled airline accidents, so that action may be taken to prevent reoccurrence. Although not required on other than air carrier aircraft, it may be possible in the future for a similar capability to be included as part of an integrated avionics systems, particularly if a low-cost crashworthy storage device can be developed.

System Monitoring. A primary use of flight instrumentation is in the monitoring of key performance parameters, such as the temperature of an engine or the amount of remaining fuel. Initially, only fundamental properties, such as quantity, temperature, or pressure could be monitored; important values such as percent of engine power or true airspeed, which cannot be measured directly by a single sensor, were not available. However, with the increased reliability and decreased cost and size of electronics, it became possible to combine more than one sensor to calculate important measurements. For example, air data computers calculate and display true airspeed based on static and pitot pressures and air temperature, offering a substantial advantage over requiring a busy flight crew member to calculate it from indicated airspeed, outside temperature, and pressure altitude (which itself is generally not directly presented). The trend is toward more automation than is currently available, using inexpensive digital microprocessors which can calculate both expected and actual engine power based on a number of variables, such as density altitude, fuel flow, throttle setting, RPM, manifold pressure, etc., for a reciprocal engine, or turbine temperatures and RPM's, fuel flow, and pressures for a jet engine.

Of even more importance is the rapidly developing trend toward using the ability of computers and electronics to monitor these readings, and sound an alarm if there is a problem. The two simplest methods are to establish an upper and lower bound for each reading, and indicate when the bound has been exceeded, or to compare two different readings and indicate when they differ by a set amount. The latter approach is used extensively on transport category airplanes, where separate systems exist or both the captain and the first officer, and a monitor assures their consistency.

A difficulty with this type of monitoring is that the problem may be caused by a sensor failure, rather than an actual system failure, so that the resulting alarm is in error. Also, what is a problem in one phase of flight may not be in another. For instance, a descent which will cause contact with the ground should be detected and the crew alerted during the takeoff and enroute segments of the flight, but not during the final phase of the landing. If the monitor is not intelligent enough to recognize these two cases, erroneous alarms will result, with the flight crew becoming conditioned to ignore them (even when they are actually valid). Cockpit recordings from many accidents show that an automatic alarm (such as an altitude alert) was given, but was ignored by the crew. Since this desirable trend toward automatic monitoring and reporting is so strong, and the historical precedents indicate that a relatively few false indicators will cause future failure reports to be ignored, there is already a need for techniques in redundancy management which will make monitoring effective.

Work is currently progressing in many areas, including the application of artificial intelligence concepts to cockpit automation [3]. When a sensor indicates a value which is out of range, special procedures, using a description of the interdependencies of the various system components, verifies whether the problem exists in the system or in the sensor. Other parts of the system perform a flight following function, to determine the context in which an unusual reading exists. From this, the criticality of the problem can be assessed, and action taken based on this. For example, if there is sufficient time, an interactive dialog can be initiated with the crew to determine the optimal course of action. However, if time is critical, a vocal alarm can be given, directing the pilot to perform the necessary maneuvers. The system acts as an ever-vigilant member of the cockpit crew, able to offer timely advice or orders, but not flying the plane itself.

Operational Cost Reduction

Flight instrumentation has reduced the cost of operating aircraft in a number of ways. Obviously, the increase in safety and the consequent reduction in accidents reduces costs. Increases in capability, discussed in the following section, also favorably affect costs. However, the major contributions are in the reduction of required crew, increased efficiency, and assistance in maintenance.

Crew Reduction. The development of the autopilot, with its ability to accurately control the aircraft during many maneuvers, relieves the flight crew of simple procedures to better concentrate on important flight considerations. This eliminates the necessity of an increased crew size, with resulting communications and logistics problems, as the workload in high stress situations increases. In the air taxi area, autopilot systems can is some cases be substituted for a second crew member (FAR 135.77), reducing expense and allowing an extra revenue passenger in a small aircraft. Recently, a single pilot version of the Cessna Citation business jet was certificated by the FAA, primarily by the substitution of flight control electronics for the previously required second pilot.

Automatic global navigation systems, such as Omega, inertial, and doppler radar, have virtually eliminated the task of the navigator by providing instantaneous position information far more accurately than previous manual systems. In addition, these navigation systems, in addition to RNAV, allow more direct routing of the flight, reducing its cost by decreasing the route length. Proposed advances, such as four-dimensional RNAV, promise to eliminate most holding maneuvers.

Energy Management. Because of the substantially increased portion which fuel costs now contribute to total aircraft operating expenses, energy management has become an extremely important concern for aircraft operators. While pilots were previously instructed to fly departures, enroute segments, and arrivals according to handbook values, they were often too busy to do this precisely. However, a substantial reduction in operational costs can be achieved through careful management of the engines and the operating altitude of the aircraft. Computers monitoring an aircraft's position and altitude, which contain information about winds aloft, and know the current fuel flow and the values to be expected at other power settings can be used to calculate whether the additional fuel expended in reaching a higher altitude will be rewarded by a comparable decrease in fuel consumption. Optimal climb rates and descent points can be calculated, further improving efficiency. Not only do these techniques reduce the fuel used, but since less fuel must be carried during the flight, and hence less weight, additional savings are possible. It has been estimated that only 70 to 80 percent of tankered fuel is constructively used -- the other 20 to 30 percent is necessary to carry the additional weight of the tankered fuel. The trend is to directly couple this information to the fligth control system (autopilot/autothrottle) as an addition input to allow it to optimize aircraft performance.

Autopilot systems can also be used to provide artificial stability to an aircraft, reducing the need for dynamic stability to be designed into an airframe. This technique allows the wings and other airfoils to be attached at optimal angles, rather than at a dihedral, so that drag is reduced. This in turn reduces the power necessary in flight, and hence the fuel burned. It is only with the latest advances in digital computer reliability that such a system can be considered, since it was not possible using previous electronics to achieve the level of safety required by the FAA for such a flight critical application.

<u>Diagnostic Capabilities.</u> Significant steps have been taken in recent years to provide some self diagnosis of system failures, principally for maintenance purposes, and including in some cases pilot alerting of critical system status changes. The trend toward systems diagnosing their own difficulties, and indicating the failing components, offers substantial potential reduc[illegible]ons in thest to repair avionics systems. The diagnostics can indicate the board

containing the failure, which can then be replaced by the service engineer. The failed board is returned to the manufacturer, where it is tested further using special automated fault detection equipment. While it may seem expensive to replace an entire board when only a simple part such as a diode has failed, as labor costs for field service personnel increase it is actually less expensive. This diagnostic capability can be extended to automatically diagnose and record important information whenever any problem, either permanent or transient, is encountered. This data can be summarized to indicate potential problems which can then be serviced before they cause difficulties.

Expanded Capabilities

Many of the flight operations taken for granted today would not be possible without the contribution of flight instrumentation. The greatest advances have been in the area of all weather operations (IFR flight). Without visual reference or proper instrumentation, no pilot is able to control his aircraft successfully. It is only a question of time before the airplane enters an unusual attitude, generally a spiral of increasing speed which may only stop when the ground is reached. Flight instrumentation, even as basic as an airspeed indicator, an altimeter, and a rate of turn indicator, can allow a trained pilot to retain control throughout a flight. Autopilot systems, previously discussed, and sophisticated instrumentation such as flight directors, simplifies the instrument flight task.

Electronic navigation systems provide the second component for all weather operation -- the ability to determine one's position without seeing landmarks on the ground. A variety of systems have been proposed and implemented for this purpose, including VOR, ILS, NDB, and Omega. Just as it is hard to find an airplane which does not have two-way radio capability, few are not equipped with VOR navigation equipment. Even student pilots flying in VFR weather conditions, use VOR as a primary navigation method, resulting in fewer lost airplanes.

While advanced navigation systems have always been used by the airlines to allow the meeting of schedules during all but the worst weather, an increasing number of general aviation airplanes are fully instrument equipped, including full ILS capability, area navigation, and flight directors.

Approach guidance aides, such as the current instrument landing system (ILS) and the proposed microwave landing system (MLS) can be extended to provide "blind" landing capabilities. Category I (200 foot decision height and 2400 foot visibility) and Category II (100/1200) are common, and Category IIIA (automatic landing, visual rollout) approaches have been tested at a number of airports and will be shortly approved as standard operations. Most wide-body transport aircraft are equipped with the special autopilot systems to allow fully automated landings. Recent advances in flight control and stability augmentation systems have allowed helicopters, which are basically unstable, to be certified by the FAA for single pilot IFR operation. This will have substantial impact in the future in their use for business transportation and for operation between offshore oil platforms and the mainland.

Other systems, such as weather radar or other thunderstorm detection equipment and deicing equipment, allow safe operation in potentially dangerous situations. In both transport category and general aviation, nothing has increased the utility of the basic aircraft as much as improved flight instrumentation and control systems.

COST/PERFORMANCE CONSIDERATIONS

In the previous section, the contributions of flight instrumentation and control systems to aviation were discussed. However, the specific tradeoffs between cost and performance among various systems were ignored. These factors must be considered if the worth of a particular system is to be determined, or if similar systems are to be compared. The obvious method for determining the worth of anything, including flight instrumentation systems, it to add up the value of its benefits and subtract its cost.

While this may seem like an easy calculation, in practice the determination of the value of various cost and benefit components may prove to be difficult. This is especially true when second order costs and benefits are considered, as must be the case when comparing two closely matched systems (such as different mechanizations of the same concept, like MLS). Costs are perhaps the easiest to compute, with "life-cycle costs" calculations a good example of the pertinent techniques. A major difficulty is predicting the cost of goods and services in the future, which requires an estimate of future inflation rates and other variables.

Substantially more difficult to compute, and therefore sometimes ignored in preliminary comparisons of competitive systems, are the benefits derived by the use of a particular system. When comparing systems of comparable functional specifications, such as various landing guidance systems, it is reasonable to ignore the consideration of benefits, since they are essentially equal for the various systems. However, when computing the absolute worth of a system, comparing systems which are not functionally identical (such as VOR/DME to Omega), or determining the value of a system improvement (such as from Category I to Category II), the resulting benefits derived become a substantial consideration.

If estimating costs is difficult, estimating benefits can be virtually impossible. They are based on a number of intangibles, such as the price of safety. While absolute safety is a desirable goal, it is not possible to achieve at finite cost. Therefore, what level is satisfactory? In the equation for the worth of a system, what value should be placed on a life? These questions, and their effect on the cost/benefit equations for a system, will be considered in more detail later.

Other benefits do not involve such philosophical questions, but can be difficult to determine. For instance, the benefits of an approach guidance system which has Category II minima, as opposed to Category I minima, is a function of the number of approaches expected during the time when the weather is at or better than Category II minima, but less that Category I. In addition, costs of diversion to alternate airports, missed approaches, and holding until better weather must be considered, along with a potential decrease in airport capacity if higher separation standards may be used. The first factors can be predicted from past statistical weather information although it may require extensive searching, while the second group could be derived from simulation models of representative terminal areas which consider the procedures of various air carriers.

The following discussion covers most of the cost factors and benefits which must be considered to properly evaluate the worth of a given system. In many instances, it indicates representative values for given systems. Initially, the cost considerations, which are probably the most important factors in comparing systems, are discussed.

Station Costs

These costs are for all components of a system which are not installed on the aircraft utilizing the system. For self-contained navigation systems, such as inertial, for flight control systems, or for many other flight instrumentation and control systems, there is no station costs. However, for radionavigation systems, such as VORTAC, LORAN, Omega, ILS, and MLS, the station costs can be a substantial factor. Previously these costs were referred to as ground station costs, but with the advent of earth satellite stations for use in navigation and communication the more general term is necessary. These costs can be divided into two classifications -- purchase and siting, and operation and maintenance.

Purchase and Siting Costs. This is simply the cost of all the station equipment and the cost of putting it in place. For ground-based stations, this includes the purchase cost for land, any work necessary to adapt it for use by the system (such as earthmoving to change the topography or removal of existing structures), and construction costs for necessary structures and for electrical service. For satellite stations, the siting costs are the costs of launching it into orbit.

The purchase price for the equipment of an existing system can be determined by either using the past capital investment (which will generally also include siting costs, projected to a common date to properly compensate for inflationary effects), or the cost of replacement for the system. For example, the total capital investment for the VORTAC system at the present time is about $250 million [13], while the cost of replacement of the VOR portion for the existing 921 ground stations would be about $23 million (based on a cost of $14,000 for a single transmitter system and $25,000 for a dual installation; figures provided by FAA AAF-410). The equipment for the eight Omega stations was purchased for approximately $1.5 million per station, and would cost about $3 million at the present time (figures provided by Navy PME-119).

For a proposed system, the equipment cost can be predicted by determining the number of stations required, and multiplying by the average station cost. This determination requires decisions regarding service areas (such as low altitude coverage in mountainous terrain for a line of sight system or which runways will receive a class of approach aid), acceptable noise and geometric error, and number of expected users if the system requires active communications between the user and the station (such as radar and DME transponders). Examples of this calculation can be found as part of the justification of virtually all proposed systems.

Operating and Maintenance Expenses. This includes all expenses to support the normal operation of the station, including power costs, operating personnel, maintenance personnel, replacement parts, etc. For existing systems, appropriate data is generally available, although possibly not in a convenient form. For VORTAC, the costs last year were approximately \$36.3 million, with \$21.6 million spent on VOR and the remaining \$14.7 million for TACAN (figures from FAA AAF-250). When the proposed modernization program is completed, replacing the obsolete tube equipment with solid state units, this cost should be substantially decreased due both to a lower power consumption and less preventive and unscheduled maintenance. For Omega, the yearly operating and maintenance costs, paid by a number of countries, is about \$6.75 to \$7.5 million (figures provided by Navy PME-119).

If the station is not capable of automatic operation, but requires operational personnel (such as radar or direction finding stations), their salaries and benefits must be included as a part of the operating expenses. Because of this, most new systems require only minimal operator intervention (for example, ILS and MLS replacing PAR/GCA).

It is probable that over the life of a system, the operating and maintenance costs will exceed the capital investment for the system, making careful estimation and calculation of this cost factor especially important when deciding the worth of a new or existing system.

Development and Certification Costs

While these costs can be substantial, they are generally not considered directly when computing the total cost of a system. This is because, especially for new avionics systems, the initial system development costs are included in government research and development programs. For the Omega navigation system and the proposed microwave landing system (MLS) over \$100 million has been spent for research and development. Since it can represent a substantial investment, it should be considered when a new system is being contemplated, to determine if the system's benefits are worth its cost.

The costs of developing individual receivers and other system components are generally included as part of the purchase price. Also amortized over the number of units produced is the cost of testing and certificating a unit. For complicated systems, this can be of considerable expense even though there is no charge made by the Federal Aviation Administration for this work. The high cost is primarily due to the extensive testing required to assure that the equipment will perform within specification during all expected operating conditions. The use of standardized computer equipment, instead of custom designed logic, will both aid and complicate this procedure. Since standard equipment is being used, particularly integrated circuits which have received military qualification, less time is necessary than for specially designed circuits.

However, this will shift the emphasis in testing and verification from hardware to software. Although for simple programs it is possible to exhaustively test to determine if its operation is correct for all valid inputs, and that a reasonable action is taken if an invalid input occurs, the state of the art in software design and verification techniques has not produced any single method of checking large, complex programs such as found in aviation applications. While work is in progress at a number of institutions, including the MUST project at NASA Langely Research Center, currently the certification of software is a costly and complex task.

Training Costs

The expense of retraining service personnel and pilots for a new system must be evaluated when determining the final cost of any system. For service personnel, this expense could be minimal, if the system is closely related to existing systems (such as the interim standard MLS to conventional ILS, where the primary difference is a frequency converter front end providing a shift from microwave frequencies to those used by ILS). However, if it is based on substantially different technology, such as a shift from analog processing to digital, a substantial cost of training may result.

A similar situation exists for pilots, except that (if a system is adopted as a standard) the training is only to re-educate existing pilots, with new pilots being educated in the use of the system as part of their training. For example, when the LF ranges were being replaced with the VOR system, it was necessary to retrain existing pilots, at an additional cost. However, new pilots received instruction in VOR as a normal part of their flight training. To a limited extent, this is now occurring as some instrument flight training programs offer an introduction to RNAV.

To minimize the retraining expenses, one of two different approaches can be used. First, the user interface to the system can be simplified so that very little training is necessary to use the system. DME functions this way, since it only requires the selection of a frequency matching that of the desired VOR to get a distance from that VOR. In fact, many DME systems automatically tune to the appropriate frequency using information from the VOR receiver. All that is necessary is for the user to learn to visualize distances on the map being used.

The second approach is to design the pilot interface so that it corresponds directly to an existing system. For example, RNAV is designed so that waypoints act as if they were standard VOR's. Many Omega and VLF navigation systems are designed so that they look and operate like existing inertial navigation systems. Using computer technology and generalized displays, it may be possible to have two different modes of operation for new systems -- compatibility, where it functions like an existing system like VOR or ILS, and native, where all the new potential can be utilitzed, although requiring additional training.

Airborne Costs

As with the station costs, airborne costs can be divided into two categories -- purchase and installation, and operating and maintenance.

__Purchase and Installation.__ This is the cost of equipping the desired aircraft with appropriate equipment, and includes both the purchase price of the equipment and the expense of installing it. Generally, both these costs can be estimated very closely, giving a value which when added to the cost of purchase and siting for the stations forms the start-up costs for a new system. This value is the one often used when comparing two different system, although during the life of a system it may be only a fraction of the total costs.

For VOR receivers, the cost of a receiver ranges from approximately $1,000 for a simple general aviation unit to approximately $10,000 for a unit meeting ARINC standards and incorporating sophisticated self-test systems. However, with the introduction of low cost, high reliability integrated circuits, even low cost receiver systems now are being designed and certified to the standards (TSO's) originally for air carrier units. Currently Omega receivers, and other VLF receiver systems, cost from $25,000 to $45,000 per unit. The Air Force has initiated a procurement of 800 units at a cost of about $12,000 per unit [13]. Finally, research supported by NASA Langely Research Center at Ohio University has produced prototypes of an Omega receiver which could be constructed for around $1,000 [2].

This capital investment contributes to a substantial inertia in the adoption of new and replacement systems, particularly in the areas of communications and navigation. There is a substantial investment in both airborne and ground equipment for current systems such as VOR (as much as $800 million for airborne receivers alone [13]), and a decision which would obsolete this equipment in a short time would be highly unpopular. Only recently did the FAA revise its control tower voice communications frequency plan to drop support for aircraft with tunable receivers but only able to transmit on about a dozen frequencies. The system now supports 90 channel communications equipment (0.1 Mhz spacing over 9 Mhz) for tower communications, while currently available communications equipment is capable of 720 channels (0.025 Mhz spacing over 18 Mhz). The extra capacity is used only for IFR communications, where it is assumed the aircraft is better equipped, with the latest addition of 360 channels for high altitude communications at the present time.

A growing expense is the cost of installation of equipment, especially if it is not done during the manufacture of the aircraft. Currently, an estimate of the cost of installing new radio equipment in a general aviation airplane is approximately 15 percent of the total cost of the system. It is lowest when the equipment is totally separate from other previously installed equipment or when provisions have been made for its installation, and is highest for systems which depend heavily on other systems or require special flight test or certification procedures (such as RNAV added to existing VOR and DME equipment and being approved for IFR operations). Integrated avionics systems, using a buss structure and a common packaging technique, such as proposed in the advanced integrated avionics system under development by NASA Ames Research Center, will substantially reduce this expense, possibly allowing an experienced user to install a new system function simply by plugging in a new circuit card.

Maintenance. Cost of servicing and repairing flight instrumentation and control systems often will exceed the cost of purchase and installation over the life of a piece of equipment. The hourly cost of maintenance can be calculated using the formula:

$$\text{Cost} = \text{PM} + (\text{MTTR} \times \text{Labor} + \text{Parts}) / \text{MTBF}$$

where PM is the hourly expense for preventive maintenance, MTTR is the mean time to repair a failure, Labor is the prevailing labor rate for skilled personnel, Parts is the average cost for parts necessary in the repair, and MTBF is the mean time between failures. For a single VOR receiver, this cost is approximately $0.10 per hour of flight (estimate from Tom Ellison of United Airlines).

Years ago, the cost of labor was low, while the cost of replacement parts was high, making the hourly maintenance cost primarily a function of the cost of the parts expected to fail (generally vacuum tubes). However, this is no long true. The revolution in electronics starting with the transistor and continuing to today's integrated circuits has reduced the cost of failing parts to a small fraction of the total cost of repairs. In many cases, the failure will be in a part worth less than $1, while the labor for an hour of service personnel time may cost at least $20. Therefore, the cost of parts may be ignored in the equation, giving a new equation:

$$\text{Cost} = (\text{MTTR} \times \text{Labor}) / \text{MTBF}$$

Since the cost of labor will be increasing with time (the $20 per hour expense is low when compared to the rates for similarly skilled service personnel, such as computer service engineers at over $40 per hour), the only way to reduce the final cost is to either increase the MTBF or reduce the MTTR.

Current electronics technology, particularly integrated circuits, currently produces high MTBF figures after an initial "infant mortality" time, generally less than 100 hours. Most failures can be traced to operation outside the desired enviromental conditions (high temperature or humidity), thermal or mechanical shock, or failure of mechanical connectors or controls. New digitally controlled systems reduce the last consideration, while proper operating techniques minimize the first two. <u>Substantially increased MTBF's should not be expected in the future.</u>

<u>On the other hand, improvements can be made to the MTTR, especially for complex systems.</u> Additional circuits can be added to equipment to aid in the testing and isolation of difficulties. The use of digital computers allows for special diagnostic programs which can log any error, analyze the total system either when an error occurs or when requested by the user, or determine the location of the failing component. Instead of highly paid service personnel spending time determining the cause of a problem, the system can direct less skilled personnel to remove and replace a given circuit card, reducing both the cost of labor and the MTTR. The bad card can then be tested at a special service depot using automated board testers, to determine the actual cause of the failure. This procedure is especially attractive when the problem occurs only intermittently, since the computer can make a diagnosis at the time of failure, rather than later when it might not occur.

The advances in system design caused by solid state technology, improved packaging, integrated systems, electronic switching and control rather than mechanical, and self-testing and automatic diagnostic programs can make the operating and maintenance costs much less a factor than was the case with vacuum tube equipment. It may be that this cost will be substantially replaced by a depreciation cost for the equipment, assuming that it will be totally replaced within a given number of years regardless of its condition.

Cost Reductions

Obviously, a major benefit of one flight instrumentation and control system over another is a reduction in costs, either in the purchase and installation of the units, their operating and maintenance costs, or their reduction or elimination of another expense. In the latter category is the reduction of the required flight crew, use of less fuel, and other efficiency increases. Determining the benefits derived from a specific flight instrumentation and control system requires the development of models of aircraft operations for specific categories of users. For example, it may be possible to generalize the effects of a system for all transport category jet aircraft, all non-jet transport category aircraft, general aviation aircraft used in business transportation, and general aviation aircraft used in non-business applications (such as training). The calculated benefits from each of these models can then be multiplied by the number of flight hours or operations for each of the categories to give a final estimate for the benefits. It should be noted that the various categories are based on equivalence classes for a particular system and operation, and may change drastically when another system is analyzed. For example, if the system being studied has as a potential benefit the elimination of a separate flight navigator, then the two classes which need to be studied are those aircraft which utilize a navigator, and those which don't. When the benefits from the first class have been calculated and multiplied by the expected number of flight hours, the total benefit will result. It makes no difference how the classes are formed, since only the "bottom line" figure is of interest (although the intermediate figures may give information for the later optimization of the system benefits).

Flight Crew Reduction. While this is an important benefit of the advances in flight instrumentation and control systems, it typically does not occur as an actual reduction in a flight crew size. Most general aviation operations are single pilot flights, so obviously no reduction is possible. The size of air carrier flight crews is determined jointly by federal regulation, company policy, and union contracts. Only in limited areas, such as that of the air taxi operations discussed previously, is an actual reduction in crew size possible.

However, present flight crews are now capable of more complicated operation without size increases. General aviation airplanes regularly fly in IFR weather conditions with only a single pilot; helicopter IFR operations with only a single pilot are now possible, due primarily to advanced stability augmentation systems. Therefore, the models developed should determine the number of personnel required for equivalent workloads on the remaining crew with and without a given system. In the case of helicopter IFR, this would be equal to a savings of one crew member for each flight in which IFR conditions are possible. Multiplying the rate of pay by the number of flight hours gives the expected benefit for the use of advanced systems.

Energy Savings. Again, a model for the expected operations is required, which gives a percentage savings in fuel usage by the adoption of a given system. This can then be multiplied by the fuel costs for a given segment of aviation, and the result combined with those from other segments to give a final expected benefit. Since fuel costs have been increasing rapidly in the past years, care must be taken when this calculation is made to include both the effects of normal inflation and of increasing fuel costs due to reductions in supplies. In the future, this may prove to be the greatest benefit of advanced flight instrumentation and control systems.

Weight Reduction. An aircraft is designed to carry a given payload, which can either be revenue producing (cargo or passengers) or dead weight (the airframe, required crew, and all installed systems). It is estimated that the cost of transporting dead weight is about $40 per year per kilogram, although the actual cost is of course a function of the particular type of aircraft, fuel costs, and flight hours. However, systems which result in a lighter weight than other systems can produce substantial benefits when all aircraft are considered.

However, the greatest potential for weight reduction comes from flight controls which reduce the need for complex and heavy structures in the airframe. For example, the advanced lift distribution control system used as a modification to the Air Force's C5 transport airplane has produced an increased payload, and, of possibly more benefit, an increase of the expected life of the wing to 30,000 flight hours (figures provided by Lockheed C5 project management).

Expanded Capabilities

The important contributions of advanced flight instrumentation and control systems towards the development of advanced capabilities, such as high altitude or all weather operations, were discussed previously. It is in this area that the most striking benefits of advanced systems have occurred, particularly when compared to the abilities of the basic aircraft without these systems. Again, determining the benefits requires the definition of equivalence classes for a particular improvement, and the development of an appropriate model for each class. For example, to calculate the effect of ILS Category II operations over Category I, weather information and flight schedules must be compared to determine how many flights are affected by the difference in weather minima for an approach. Company policy regarding dispatch into weather which may be below minima, expected holding times, and distances to alternate airports must be considered. When these are appropriately combined, an approximate value for the benefits of lower IFR approach minima results. Calculations similar to these have been carried out by the planning staffs of most major airlines to determine if the additional cost of an advanced system is offset by its benefits. For example, United Airlines determined a number of years ago that Category II had sufficient benefits to warrant the expense, but currently does not feel that the benefits of Category III outweigh its cost. In addition, it has been found that the use of inertial navigation equipment has sufficient benefits only on those flights going to or from Hawaii; the equipment is installed before the flight leaves the mainland, and is removed upon its return, to reduce the maintenance and dead weight expenses (information from United Airlines).

Increased Safety

It is more difficult to arrive at a value for this benefit than for all others, although it is probably the most important benefit of all. For those improvements which are not directly related to flight safety, such as improved navigation systems, it is reasonable to assume that the flights will be operated within appropriate conditions, such as instrument approaches only to those minima allowed for a given system or flight near a dangerous area (thunderstorms, high terrain) only if adequate guidance is available. In this case, there is no change in the safety of the flight, only in whether it can be conducted, and the model used for assessing expanded capabilities can be used.

For equipment specifically installed to enhance safety, such a ground proximity warning systems (GPWS) or emergency locator transmitters (ELT), the determination of the benefits is not as easy. Probably the best method is to determine the occurrence rate for the class of accidents or incidents which the system prevents, multiply it by the expected improvement, and by the projected number of flight hours. This will give a figure of the number of accidents or incidents which will be prevented by the system, which can then be multiplied by the average past cost of similar accidents to arrive at a value.

A major difficulty exists in this procedure. It requires a prediction of the amount of court awards in the case of injuries or deaths, a very difficult task (if even possible). However, systems designed specifically to improve safety, with no other improvements in capabilities, typically have been instituted not as the result of cost/benefit planning, but as a response by Congress or regulators to one or more accidents. Therefore, a comparative analysis of these systems may not be as important as with other systems.

CONCLUSIONS

As can be seen from historical review, flight instrumentation and control systems have played a vital part in changing aviation from an unsafe novelty to a reliable form of commercial transportation. One of the most important contribution is the ability to safely and reliably navigate between distant points and within congested terminal areas. Central to this capability is the ability of flight instrumentation, in conjunction with ground-based counterparts, to more accurately specify an aircraft's coordinates in space. Improvement in this accuracy is still required to meet foreseeable demands, with the technological opportunities barely keeping pace with requirements (such as with MLS and GPS/Navstar).

It is also clear from this review that the environment in which the commercial aircraft operates is transitioning from one of extreme growth to moderate growth with a shifting emphasis on improved efficiency. Today's technology driver is economics, and the two greatest economic drivers today are energy and labor costs. The avionics contributions of the future must address these two areas directly.

In the case of energy cost reduction, the technical trade-offs are between aerodynamically and structurally stable aircraft designs, with their inherent safety, on one hand, and efficiency on the other. The degree of this trade-off that can be realized will be strictly a function of future avionics capability. In arriving at a suitable approach to providing this opportunity for high efficiency, the avionics systems must not, throught their costs, negate the value of the improved efficiency. Therefore, an emphasis in future research and development activities must be toward reducing purchase and maintenance costs.

Finally, there is the question of reliability. If a significant improvement in efficiency is obtained via the active application of new avionics systems, they will become flight critical systems subject to much more stringent requirements than present day avionics, since they will be as important to the safety of flight as the wings and engines.

REFERENCES

[1] Black, Archibald. The Story of Flying, Whittlesey House (McGraw-Hill), London, 1940.

[2] Burhans, R W. "Phase-difference method offers low-cost navigation receivers", Electronics, September 5, 1974.

[3] Chien, R T. Artificial Intelligence and Human Error Prevention -- A Computer Aided Decision Making Approach, Report T-53, University of Illinois Coordinated Science Laboratory, February, 1978.

[4] DOT-FAA Air Traffic Control, Handbook 7110.65A, January 1, 1978.

[5] Drouihet, P R Jr. "ATC Application of the DABS Data Link", AIAA Paper 77-1510.

[6] Eleccion, Marce. "The promise of air safety", IEEE Spectrum, July 1975.

[7] Grover, J H H. Radio Aids to Air Navigation, Philosophical Library, New York, 1957.

[8] Kayton, M and W R Fried (ed). Avionics Navigation Systems, John Wiley & Sons, 1969.

[9] Litchford, G B. "Avoiding midair collisions", IEEE Spectrum, September 1975.

[10] Molloy, E (ed). Aeroplane Instruments (Part 1), Chemical Publishing, New York, 1940.

[11] Molloy, E (ed). Instruments (Part 2), George Newnes Ltd, London.

[12] Stewart, C J. Aircraft Instruments, Chapman & Hall, London, 1930.

[13] Yulo, Carlo. "An Overview of En Route Radio Navigation Services for Civil Aviation", AIAA Paper 77-1509